

# A Convex Programming-based Algorithm for Mean Payoff Stochastic Games with Perfect Information\*

Endre Boros<sup>†</sup>   Khaled Elbassioni<sup>‡</sup>   Vladimir Gurvich<sup>§</sup>   Kazuhisa Makino<sup>¶</sup>

## Abstract

We consider two-person zero-sum stochastic mean payoff games with perfect information, or BWR-games, given by a digraph  $G = (V, E)$ , with local rewards  $r : E \rightarrow \mathbb{Z}$ , and three types of positions: black  $V_B$ , white  $V_W$ , and random  $V_R$  forming a partition of  $V$ . It is a long-standing open question whether a polynomial time algorithm for BWR-games exists, even when  $|V_R| = 0$ . In fact, a pseudo-polynomial algorithm for BWR-games would already imply their polynomial solvability. In this short note, we show that BWR-games can be solved via convex programming in pseudo-polynomial time if the number of random positions is a constant.

## 1 Introduction

We consider two-person zero-sum stochastic games with perfect information and mean payoff: Let  $G = (V, E)$  be a digraph whose vertex-set  $V$  is partitioned into three subsets  $V = V_B \cup V_W \cup V_R$  that correspond to black, white, and random positions, controlled respectively, by two players, MIN - the *minimizer* and MAX - the *maximizer*, and by nature. We also fix a *local reward* function  $r : E \rightarrow \mathbb{Z}$ , and probabilities  $p(v, u) > 0$  for all arcs  $(v, u)$  going out of  $v \in V_R$ . We assume that  $\sum_{u|(v,u) \in E} p(v, u) = 1$ , for all  $v \in V_R$ . Vertices  $v \in V$  and arcs  $e \in E$  are called *positions* and *moves*, respectively. The game begins at time  $t = 0$  in the initial position  $s_0 = v_0$ . In a general step, in time  $t$ , we are at position  $s_t \in V$ . The player who controls  $s_t$  chooses an outgoing arc  $e_{t+1} = (s_t, v) \in E$ , and the game moves to position  $s_{t+1} = v$ . If  $s_t \in V_R$  then an outgoing arc is chosen with the given probability  $p(s_t, s_{t+1})$ . We assume that every vertex in  $G$  has an outgoing arc. In general, the strategy of the player is a policy by which (s)he chooses the outgoing arcs from the vertices (s)he controls. This policy may involve the knowledge of the previous steps as well as probabilistic decisions. We call a strategy *stationary* if it does not depend on the history and *pure* if it does not involve probabilistic decisions. For this type of games, it will be enough to consider only such strategies, since these games are known to be (polynomially) equivalent [BEGM13a] to the perfect information stochastic games considered by Gillette [Gil57, LL69].

In the course of this game players and nature generate an infinite sequence of edges  $\mathbf{p} = (e_1, e_2, \dots)$  (a *play*) and the corresponding real sequence  $r(\mathbf{p}) = (r(e_1), r(e_2), \dots)$  of local rewards. There is a global payoff function  $\phi$  that maps any local reward sequence to a real number, and it is assumed that MIN pays MAX the amount  $\phi(r(\mathbf{p}))$  resulting from the play. Naturally, MAX's aim is to create a play which maximizes  $\phi(r(\mathbf{p}))$ , while MIN tries to minimize it. (Let us note that the local reward function  $r : E \rightarrow \mathbb{R}$  may have negative values, and  $\phi(r(\mathbf{p}))$  may also be negative, in

\*Part of this research was done at the Mathematisches Forschungsinstitut Oberwolfach during a stay within the "Research in Pairs" Program from July 26, 2015-August 15, 2015. The research of the third author has been partially funded by the Russian Academic Excellence Project '5-100'. The fourth author was partially supported by KAKENHI Grant Numbers 24106002 and 26280001.

<sup>†</sup>MSIS Department and RUTCOR, Rutgers University, 100 Rockafellar Road, Livingston Campus Piscataway, NJ 08854, USA; (Endre.Boros@rutgers.edu)

<sup>‡</sup>Masdar Institute of Science and Technology, P.O.Box 54224, Abu Dhabi, UAE; (kelbassioni@masdar.ac.ae)

<sup>§</sup>MSIS Department and RUTCOR, Rutgers University, 100 Rockafellar Road, Livingston Campus Piscataway, NJ 08854, USA; and National Research University, Higher School of Economics, Moscow, Russia; (Vladimir.Gurvich@gmail.com)

<sup>¶</sup>Research Institute for Mathematical Sciences (RIMS) Kyoto University, Kyoto 606-8502, Japan; (makino@kurims.kyoto-u.ac.jp)

which case MAX has to pay MIN  $-\phi(r(\mathbf{p}))$ . Let us also note that  $r(\mathbf{p})$  is a random variable since random transitions occur at positions in  $V_R$ .) Here  $\phi$  stands for the *limiting mean payoff*

$$\phi(r(\mathbf{p})) = \liminf_{T \rightarrow \infty} \frac{\sum_{i=1}^T \mathbb{E}[r(e_i)]}{T}, \quad (1)$$

where  $\mathbb{E}[r(e_i)]$  is the expected reward incurred at step  $i$  of the play.

As usual, a pair of (not necessarily pure or stationary) strategies is a *saddle point* (or *equilibrium*) if neither of the players can improve individually by changing her/his strategy. The corresponding  $\phi(r(\mathbf{p}))$  is the value  $\mu_{\mathcal{G}}(v_0)$  of the game with respect to initial position  $v_0$ . Such a pair of strategies are called *optimal*; furthermore, it is called *uniformly optimal* if it provides the value of the game for any initial position. It is known [Gil57, LL69] that every such game has a pair of uniformly optimal pure stationary strategies. A BWR-game is said to be *ergodic* if  $\mu_{\mathcal{G}}(v) = \mu$  for all  $v \in V$ , that is, the value is the same from each initial position.

This class of *BWR-games* was introduced in [GKK88]; see also [CH08]. The special case when  $V_R = \emptyset$ , *BW-games*, is also known as *cyclic* games. They were introduced for the complete bipartite digraphs in [Mou76b, Mou76a], for all (not necessarily complete) bipartite digraphs in [EM79], and for arbitrary digraphs<sup>1</sup> in [GKK88]. A more special case was considered extensively in the literature under the name of *parity games* [BV01a, BV01b, CJH04, Hal07, Jur98, JPZ06], and later generalized also to include random positions in [CH08]. A BWR-game is reduced to a *minimum mean cycle problem* in case  $V_W = V_R = \emptyset$ , see, e.g., [Kar78]. If one of the sets  $V_B$  or  $V_W$  is empty, we obtain a *Markov decision process* (MDP), which can be expressed as a linear program; see, e.g., [MO70]. Finally, if both are empty,  $V_B = V_W = \emptyset$ , we get a *weighted Markov chain*. For BW-games several pseudo-polynomial and subexponential algorithms are known [GKK88, KL93, ZP96, Pis99, BV01a, BV01b, HBV04, BV05, BV07, Hal07, Sch09, Vor08]; see also [JPZ06] for parity games. Besides their many applications (see e.g. [Lit96, Jur00]), all these games are of interest to Complexity Theory: It is known [KL93, ZP96] that the decision problem “whether the value of a BW-game is positive” is in the intersection of NP and co-NP. Yet, no polynomial algorithm is known for these games, see e.g., the survey by Vorobyov [Vor08]. A similar complexity claim can be shown to hold for BWR-games, see [AM09, BEGM13a].

## Main result

The computational complexity of stochastic games with perfect information is an outstanding open question; see, e.g., the survey [RF91]. While there are numerous pseudo-polynomial algorithms known for the BW-case, it is a challenging open question whether a pseudo-polynomial algorithm exists for BWR-games, as the existence of such an algorithm would imply also the polynomial solvability of this class of games [AM09].

In [BEGM13b, BEGM15], we gave a pseudo-polynomial algorithm for BWR-games when *the number of random positions is fixed*. In this note we show that one can obtain a similar result via convex programming, combined with some of the ideas in [BEGM13b, BEGM15].

For a BWR-game  $\mathcal{G}$  let us denote by  $n = |V_W| + |V_B| + |V_R|$  the number of positions, by  $k = |V_R|$  the number of random positions, and assume (without loss of generality) that all local rewards are non-negative integers with maximum value  $U$  and all transition probabilities are rational with common denominator  $D$ . The main result of this paper is as follows.

**Theorem 1** *A BWR-game  $\mathcal{G}$  can be solved in  $\text{poly}(n, U, D^k)$  time via convex programming.*

This theorem extends the result by Schewe [Sch09], where it was shown that solving BW-games can be reduced to solving linear programming problems with *pseudo-polynomial* bit length.

According to the results in [BEGM13b, BEGM15], to get a pseudo-polynomial algorithm for BWR-games, it is enough to have pseudo-polynomial routines for: (i) solving BW-games; (ii) solving

---

<sup>1</sup>In fact, BW-games on arbitrary digraphs can be polynomially reduced to BW-games on bipartite digraphs [BEGM13a]; moreover, the latter class can further be reduced to BW-games on complete bipartite digraphs [CHKN14].

ergodic BWR-games; and (iii) finding the top and bottom classes in a non-ergodic BWR-game (that is, the sets of positions with highest and lowest values).

There are several pseudo-polynomial algorithms for solving BW-games, e.g., [GKK88, Pis99, ZP96]. One may also use the LP-based algorithm given in [Sch09]. For (ii) we show in Section 5 how to obtain the top (resp., bottom) class in a BWR-game, and a pair of strategies solving the game induced by the top (resp., bottom) class. This also provides an algorithm for ergodic BWR-games as required in (iii).

## 2 Potential transformations and canonical forms

Given a BWR-game  $\mathcal{G} = (G, p, r)$ , let us introduce a mapping  $x : V \rightarrow \mathbb{R}$ , whose values  $x(v)$  will be called *potentials*, and define the transformed reward function  $r_x : E \rightarrow \mathbb{R}$  as:

$$r_x(v, u) = r(v, u) + x(v) - x(u), \text{ where } (v, u) \in E. \quad (2)$$

It is not difficult to verify that the obtained game  $\mathcal{G}^x$  and the original game  $\mathcal{G}$  are equivalent (see [BEGM13a]). In particular, their optimal (pure stationary) strategies coincide, and their value functions also coincide:  $\mu_{\mathcal{G}^x} = \mu_{\mathcal{G}}$ .

It is known that for BW-games there exists a potential transformation such that, in the obtained game the locally optimal strategies are globally optimal, and hence, the value and optimal strategies become obvious [GKK88]. This result was extended for the more general class of BWR-games in [BEGM13a]: in the transformed game, the equilibrium value  $\mu_{\mathcal{G}}(v) = \mu_{\mathcal{G}^x}(v)$  is given simply by the maximum local reward for  $v \in V_W$ , the minimum local reward for  $v \in V_B$ , and the average local reward for  $v \in V_R$ . In this case we say that the transformed game is in *canonical* form. To define this more formally, let us use the following notation throughout this section: Given functions  $f : E \rightarrow \mathbb{R}$  and  $g : V \rightarrow \mathbb{R}$ , we define the functions  $M[f], M[g] : V \rightarrow \mathbb{R}$ .

$$M[f](v) = \begin{cases} \max_{u|(v,u) \in E} f(v, u), & \text{for } v \in V_W, \\ \min_{u|(v,u) \in E} f(v, u), & \text{for } v \in V_B, \\ \sum_{u|(v,u) \in E} p(v, u) f(v, u), & \text{for } v \in V_R. \end{cases}$$

$$M[g](v) = \begin{cases} \max_{u|(v,u) \in E} g(u), & \text{for } v \in V_W, \\ \min_{u|(v,u) \in E} g(u), & \text{for } v \in V_B, \\ \sum_{u|(v,u) \in E} p(v, u) g(u), & \text{for } v \in V_R. \end{cases}$$

We say that a BWR-game  $\mathcal{G}$  is in (strong) canonical form if there exist vectors  $\mu, x \in \mathbb{R}^V$  such that

$$(C1) \quad \mu = M[\mu] = M[r_x] \text{ and,}$$

$$(C2) \quad \text{for every } v \in V_W \cup V_B, \text{ every move } (v, u) \in E \text{ such that } \mu(v) = r_x(v, u) \text{ must also have } \mu(v) = \mu(u), \text{ or in other words, every locally optimal move } (v, u) \text{ is globally optimal.}$$

**Theorem 2 ([BEGM13a])** *For each BWR-game  $\mathcal{G}$  there is a potential transformation  $x \in \mathbb{R}^V$  that brings  $\mathcal{G}$  to canonical form with  $\|x\|_\infty \leq L := nUk(2D)^k$ . Furthermore, in a game in canonical form we have  $\mu_{\mathcal{G}} = M[r_x]$ .*

In this paper, we will provide a convex programming formulation based on the existence of potential transformations.

We will need the following upper bound on the required accuracy.

**Lemma 1 ([BEGM10, BEGM15])** *For any position  $v$  in the top (resp., bottom) class in a BWR-game  $\mathcal{G}$ , the value  $\mu_{\mathcal{G}}(v)$  is a rational number with a denominator at most  $\sqrt{k}2^{k/2}D^{k+1}$ .*

**Lemma 2** *Consider a BWR-game  $\mathcal{G}$  and denote by  $\mathbf{1}$  the vector of all ones. Then there exists a potential vector  $x \in \mathbb{R}^V$  and  $t \in \mathbb{R}$  such that  $M[r_x] \geq t\mathbf{1}$  if and only if  $\mu_{\mathcal{G}} \geq t\mathbf{1}$ .*

**Proof** Indeed, if MAX (MIN) applies a locally optimal strategy  $s_W$  in the transformed game  $\mathcal{G}^x$  then after every own move (s)he will get (pay) at least  $t$ , while for each move of the opponent the local reward will be at least (at most)  $t$ , and finally, for each random position the expected local reward is at least  $t$ . Thus, the expected local reward  $\mathbb{E}[r_x(e_i)]$  at each step of the play is at least  $t$ . Hence, by (1), strategy  $s_W$  guarantees MAX at least  $t$  from any starting position.

The other direction follows from Theorem 2.  $\square$

A symmetric version of Lemma 2 can also be obtained by similar arguments.

**Lemma 3** *Consider a BWR-game  $\mathcal{G}$ . Then there exists a potential vector  $x \in \mathbb{R}^V$  and  $t \in \mathbb{R}$  such that  $M[r_x] \leq t\mathbf{1}$  if and only if  $\mu_{\mathcal{G}} \leq t\mathbf{1}$ .*

### 3 The convex programs

The following simple facts relate the *softmax* (resp., *softmin*) to the maximum (resp., minimum) of a set of numbers.

**Fact 1** *For any numbers  $a_1, \dots, a_n \in \mathbb{R}$  and  $b > 1$ :*

$$(i) \max_i a_i \leq \log_b \sum_i b^{a_i} \leq \max_i a_i + \log_b n;$$

$$(ii) \min_i a_i \geq -\log_b \sum_i b^{-a_i} \geq \min_i a_i - \log_b n.$$

**Proof** This follows from the fact the trivial inequalities  $b^{\max_i a_i} \leq \sum_i b^{a_i} \leq nb^{\max_i a_i}$ .  $\square$

**Fact 2** *Let  $\alpha_1, \dots, \alpha_n > 0$  be given numbers such that  $\sum_{i=1}^n \alpha_i = 1$ . Then the function  $f(x) = \prod_{i=1}^n x_i^{\alpha_i}$  is concave for  $x \geq 0$ .*

**Proof** Note that for any  $x, y \in \mathbb{R}_+^n$ , if for some  $i$ ,  $x_i = 0$  then for any  $\lambda \in [0, 1]$ ,

$$\begin{aligned} \lambda f(x) + (1 - \lambda)f(y) &= (1 - \lambda)f(y) \\ &= (1 - \lambda) \prod_{i=1}^n y_i^{\alpha_i} = \prod_{i=1}^n ((1 - \lambda)y_i)^{\alpha_i} \\ &\leq \prod_{i=1}^n (\lambda x_i + (1 - \lambda)y_i)^{\alpha_i} = f(\lambda \mathbf{x} + (1 - \lambda)y). \end{aligned}$$

Thus, it is enough to show that  $\nabla^2 f(x)$  is a negative semi-definite matrix for  $\mathbf{x} > 0$ . Note that

$$\begin{aligned} \frac{\partial f}{\partial x_i} &= \frac{\alpha_i}{x_i} f(x), \quad \text{for } i = 1, \dots, n \\ \frac{\partial^2 f}{\partial x_i^2} &= \frac{\alpha_i(\alpha_i - 1)}{x_i^2} f(x), \quad \text{for } i = 1, \dots, n \\ \frac{\partial^2 f}{\partial x_i \partial x_j} &= \frac{\alpha_i \alpha_j}{x_i x_j} f(x), \quad \text{for } i, j = 1, \dots, n, \quad i \neq j. \end{aligned}$$

Consider any  $y \in \mathbb{R}^n$ . Then

$$\begin{aligned} y^T \nabla^2 f(x) y &= \left( \sum_i \alpha_i(\alpha_i - 1) \frac{y_i^2}{x_i^2} + \sum_{i \neq j} \alpha_i \alpha_j \frac{y_i y_j}{x_i x_j} \right) f(x) \\ &= \left( \left( \sum_i \alpha_i \frac{y_i}{x_i} \right)^2 - \sum_i \alpha_i \left( \frac{y_i}{x_i} \right)^2 \right) f(x) \leq 0, \end{aligned}$$

where the last inequality follows from Jensen's inequality applied to the convex function  $f(w) = w^2$ .

□

Given  $t \in \mathbb{R}$ , let us replace the max operator in the system  $M[r_x] \geq t$  by the softmax approximation:

$$\log_b \sum_{u|(v,u) \in E} b^{r(v,u)+x(v)-x(u)} \geq t, \quad \text{for } v \in V_W, \quad (3)$$

$$r(v, u) + x(v) - x(u) \geq t, \quad \text{for } u \text{ s.t. } (v, u) \in E \text{ for } v \in V_B, \quad (4)$$

$$\sum_{u|(v,u) \in E} p(v, u)(r(v, u) + x(v) - x(u)) \geq t, \quad \text{for } v \in V_R, \quad (5)$$

where the constant  $b$  will be determined later. Defining the new variables  $y(v) := b^{-x(v)}$ , we can rewrite (3)-(5) as follows:

$$\sum_{u|(v,u) \in E} b^{r(v,u)} y(u) \geq b^t y(v), \quad \text{for } v \in V_W, \quad (6)$$

$$b^{r(v,u)} y(u) \geq b^t y(v), \quad \text{for } u \text{ s.t. } (v, u) \in E \text{ for } v \in V_B, \quad (7)$$

$$\prod_{u|(v,u) \in E} (b^{r(v,u)} y(u))^{p(v,u)} \geq b^t y(v), \quad \text{for } v \in V_R. \quad (8)$$

Note that  $y(v) > 0$  if and only if  $x(v)$  is finite. In fact, since we may assume by Theorem 2 that  $\|x\|_\infty \leq L$ , we may add also the inequalities:

$$b^{-L} \leq y(v) \leq b^L, \quad \text{for } v \in V.$$

Note that, without the lower bounds  $y(v) \geq b^{-L}$ , the system (6)-(8) is always feasible. As we shall see later, it will be necessary to test the feasibility of the system with  $y(v) > 0$  for some  $v \in V$ . For convenience, let us write more generally the following set of upper and lower bounds, where  $V' \subseteq V$  is to be chosen later:

$$0 \leq y(v) \leq b^L, \quad \text{for } v \in V, \text{ and } y(v) \geq b^{-L}, \quad \text{for } v \in V'. \quad (9)$$

Similarly, we replace the min operator in the system  $M[r_x] \leq t$  by the softmin approximation:

$$r(v, u) + x(v) - x(u) \leq t, \quad \text{for } u \text{ s.t. } (v, u) \in E \text{ for } v \in V_W, \quad (10)$$

$$-\log_b \sum_{u|(v,u) \in E} b^{-r(v,u)-x(v)+x(u)} \leq t, \quad \text{for } v \in V_B, \quad (11)$$

$$\sum_{u|(v,u) \in E} p(v, u)(r(v, u) + x(v) - x(u)) \leq t, \quad \text{for } v \in V_R, \quad (12)$$

and defining the new variables  $y(v) := b^{x(v)}$ , we can rewrite (10)-(12) as follows:

$$b^{-r(v,u)} y(u) \geq b^{-t} y(v), \quad \text{for } u \text{ s.t. } (v, u) \in E \text{ for } v \in V_W, \quad (13)$$

$$\sum_{u|(v,u) \in E} b^{-r(v,u)} y(u) \geq b^{-t} y(v), \quad \text{for } v \in V_B, \quad (14)$$

$$\prod_{u|(v,u) \in E} (b^{-r(v,u)} y(u))^{p(v,u)} \geq b^{-t} y(v), \quad \text{for } v \in V_R, \quad (15)$$

together with the lower and upper bounds:

$$0 \leq y(v) \leq b^L, \quad \text{for } v \in V, \text{ and } y(v) \geq b^{-L}, \quad \text{for } v \in V'. \quad (16)$$

## 4 Solving the convex programs

We will use the ellipsoid method [Kha80, Kha84, GLS88]. For this we need to show that the separation problem can be solved in polynomial time. For convenience, let us consider the following relaxation of the convex programs (6)-(9) and (13)-(16):

$$\sum_{u|(v,u) \in E} b^{r(v,u)} y(u) \geq b^t y(v) - \delta, \quad \text{for } v \in V_W, \quad (17)$$

$$b^{r(v,u)} y(u) \geq b^t y(v) - \delta, \quad \text{for } u \text{ s.t. } (v, u) \in E \text{ for } v \in V_B, \quad (18)$$

$$\prod_{u|(v,u) \in E} (b^{r(v,u)} y(u))^{p(v,u)} \geq b^t y(v) - \delta, \quad \text{for } v \in V_R. \quad (19)$$

$$0 \leq y(v) \leq b^L + \delta, \quad \text{for } v \in V, \text{ and } b^L y(v) \geq 1, \quad \text{for } v \in V'. \quad (20)$$

$$\sum_{u|(v,u) \in E} b^{-r(v,u)} y(u) \geq b^{-t} y(v) - \delta, \quad \text{for } v \in V_W, \quad (21)$$

$$b^{-r(v,u)} y(u) \geq b^{-t} y(v) - \delta, \quad \text{for } u \text{ s.t. } (v, u) \in E \text{ for } v \in V_B, \quad (22)$$

$$\prod_{u|(v,u) \in E} (b^{-r(v,u)} y(u))^{p(v,u)} \geq b^{-t} y(v) - \delta, \quad \text{for } v \in V_R. \quad (23)$$

$$0 \leq y(v) \leq b^L + \delta, \quad \text{for } v \in V, \text{ and } b^L y(v) \geq 1, \quad \text{for } v \in V'. \quad (24)$$

where  $\delta > 0$  is a rational number that will be chosen appropriately. Let  $K$  and  $K_\delta$  be the set of  $y \in \mathbb{R}^E$  satisfying (6)-(9) and (17)-(20), respectively. Similarly, Let  $K'$  and  $K'_\delta$  be the set of  $y \in \mathbb{R}^E$  satisfying (13)-(16) and (21)-(24), respectively.

In our application, we will set  $b = n^{4\Lambda}$  and  $t \in [0, U]$  to be a rational number with denominator  $\Lambda := \sqrt{k} 2^{k/2} D^{k+1}$ . In particular,  $b^{\pm t}$  is a rational number of bit length  $\langle b^{\pm t} \rangle = O(\Lambda \log n)$ . Also by assuming without loss of generality (by scaling  $r$  and replacing  $U$  by  $UD$ ) that  $r(v, u)$  is a multiple of  $D$ ,  $b^{\pm r(v,u)}$  is a rational number of bit length  $\langle b^{\pm r(v,u)} \rangle = O(\Lambda UD \log n)$ .

**Claim 1** For  $0 < \epsilon \leq b^{-t} \delta$  and any  $y \in K$  (resp.,  $y \in K'$ ), the box  $\{y' \in \mathbb{R}^n \mid y \leq y' \leq y + \epsilon \cdot \mathbf{1}\}$  is contained in  $K_\delta$  (resp.,  $K'_\delta$ ), where  $\mathbf{1}$  is the  $n$ -dimensional vector of all ones. In particular, if  $K \neq \emptyset$  (resp.,  $K' \neq \emptyset$ ) then  $K_\delta$  (resp.,  $K'_\delta$ ) is full-dimensional.

**Proof** We prove the statement for  $K_\delta$ ; the proof for  $K'_\delta$  is similar. Clearly,  $K \subseteq K_\delta$  for  $\delta \geq 0$ . Furthermore, for any  $y \in K$  and  $S \subseteq V$ , the vector  $y'$  obtained from  $y$  by setting  $y'(u) := y(u) + \epsilon$  for  $u \in S$  and  $y'(u) := y(u)$  for  $u \in V \setminus S$  satisfies (6)-(9). Indeed, the left-hand sides of (17)-(19) increase when  $y$  is increased in the components corresponding to  $S$ , while the right-hand sides are at most  $b^t y(v) + b^t \epsilon - \delta \leq b^t y(v)$ . Also by (9),  $y'(v) \leq y(v) + \epsilon \leq b^L + \epsilon \leq b^L + \delta$ , so  $y'$  satisfies (17)-(20).  $\square$

Now we consider the (semi-weak) separation problem for  $K_\delta$  (resp.,  $K'_\delta$ ):

Given  $\bar{y} \in \mathbb{Q}^n$  and  $0 < \delta' \in \mathbb{Q}$ , either assert that  $\bar{y} \in K_\delta$  (resp.,  $\bar{y} \in K'_\delta$ ) or find a vector  $c \in \mathbb{Q}^n$  such that  $c^T y + \delta' \geq c^T \bar{y}$  for all  $y \in K_\delta$  (resp.,  $y \in K'_\delta$ ).

**Claim 2** The separation problems for  $K_\delta$  and  $K'_\delta$  can be solved in  $\text{poly}(U, D, \Lambda, \log n, \langle \bar{y} \rangle, \langle \delta' \rangle)$  time.

**Proof** We present the proof for  $K_\delta$ ; the proof for  $K'_\delta$  is similar. Clearly, we can check in  $\text{poly}(\langle \bar{y} \rangle, \langle b^t \rangle, \langle \delta' \rangle, n)$  time if  $\bar{y}$  satisfies the linear inequalities (17), (18) and (20); if one is violated, the corresponding hyperplane defines a (exact) separator  $c \in \mathbb{Q}^n$  and we are done. Assume therefore that  $\bar{y}$  satisfies (17), (18) and (20). Let us now consider an inequality of the form (19) corresponding

to  $v \in V$  violated by  $\bar{y}$ . Let  $f(y) = \prod_{u|(v,u) \in E} (b^{r(v,u)} y(u))^{p(v,u)} - b^t y(v) := A \cdot g(y) - B \cdot y(v)$ , where  $A := \prod_{u|(v,u) \in E} b^{r(v,u)p(v,u)}$ ,  $B := b^t$  and  $g(y) := \prod_{u|(v,u) \in E} y(u)^{p(v,u)}$ . Note that  $f(\bar{y}) < -\delta$  if and only if  $A^D \cdot g(\bar{y})^D < (B \cdot \bar{y}(v) - \delta)^D$ , which can be checked in  $\text{poly}(\langle \bar{y} \rangle, \langle \delta \rangle, n, \log D)$  time, as both  $A^D$  and  $B^D$  are non-negative integers<sup>2</sup>. Without loss of generality<sup>3</sup>, we assume that for every  $u \in V$  there is exactly one edge  $(v, u) \in E$ . Then  $\nabla f(y) := A \cdot \left( \frac{p(v,u)}{y(u)} : u \in V \right) g(y) - B \mathbf{1}_v$ , where  $\mathbf{1}_v$  is the unit dimensional vector with 1 in position  $v$ . Then the inequality  $-\delta \leq f(y) \leq f(\bar{y}) + \nabla f(\bar{y})^T (y - \bar{y}) < -\delta + \nabla f(\bar{y})^T (y - \bar{y})$ , valid for all  $y \in K_\delta$  by concavity of  $f(y)$ , gives a separating inequality:

$$\nabla f(\bar{y})^T y > \nabla f(\bar{y})^T \bar{y} \quad \text{for all } y \in K_\delta. \quad (25)$$

Note that the vector  $\nabla f(\bar{y})$  can be irrational (it is irrational whenever  $g(\bar{y})$  is). We define a rational approximation  $\tilde{g}$  such that  $\tilde{g} \geq g(\bar{y}) \geq \tilde{g} - \frac{\delta'}{A}$  and  $c := A \cdot \left( \frac{p(v,u)}{\tilde{y}(u)} : u \in V \right) \tilde{g} - B \mathbf{1}_v$ . Since  $r(v, u)$  is assumed to be integer multiple of  $D$ ,  $A$  is an integer and hence  $\tilde{g}$  is a rational number of bit length  $\langle \tilde{g} \rangle = \langle A \rangle + \langle \delta' \rangle$ . It follows also that  $c$  is a rational vector of bit length  $\text{poly}(UD\Lambda \log n, \langle \tilde{y} \rangle, \langle \delta' \rangle)$ . Note that

$$c^T y - \nabla f(\bar{y})^T y = A \cdot \left( \frac{p(v,u)}{\tilde{y}(u)} : u \in V \right)^T y \cdot (\tilde{g} - g(\bar{y})) \geq 0 \quad \text{for all } y \in K_\delta, \quad (26)$$

while

$$c^T \bar{y} - \nabla f(\bar{y})^T \bar{y} = A \cdot \left( \frac{p(v,u)}{\tilde{y}(u)} : u \in V \right)^T \bar{y} \cdot (\tilde{g} - g(\bar{y})) \leq \delta' \quad (27)$$

It follows from (25), (26), and (27) that  $c^T y + \delta' \geq c^T \bar{y}$  for all  $y \in K_\delta$ . □

**Lemma 4** *Given  $t \in \mathbb{R}$  and  $\delta \in (0, 1)$  we can decide in time  $\text{poly}(n, U, D^k, \log \frac{1}{\delta})$  if the system (6)-(9) (resp., (13)-(16)) is infeasible, or find  $y(v) \in [b^{-L}, b^L + \delta]$ , for  $v \in V'$  and  $y(v) \in [0, b^L + \delta]$ , for  $v \in V \setminus V'$ , such that the left hand sides of (6)-(8) (resp., (13)-(15)) are at least  $b^t y(v) - \delta$  (resp.,  $b^{-t} y(v) - \delta$ ), for all  $v \in V$ .*

**Proof** Let us consider the equivalent system (6)-(9); the proof for (21)-(24) is similar. Given a polynomial-time algorithm for the separation problem for the convex set  $K_\delta$ , a circumscribing ball of radius  $H$  for  $K_\delta$ , and any  $\epsilon' > 0$ , the ellipsoid method terminates in  $N := O(n \log \frac{1}{\epsilon'} + n^2 |\log H|)$  calls to the separation algorithm using  $\delta' = 2^{-O(N)}$ , and either (i) finds a vector  $y \in K_\delta$ , or (ii) asserts that  $\text{vol}(K_\delta) \leq \epsilon'$ ; see, e.g., Theorem 3.2.1 in [GLS88]. In the first case, we get a vector  $y$  satisfying the conditions in the statement of the lemma. In the second case, we conclude that  $K_\delta$  and hence  $K$  is empty if  $\epsilon' < (b^{-t} \delta)^n$ . Indeed by Claim 1, if  $K \neq \emptyset$  and  $\epsilon := b^{-t} \delta$ , then  $\text{vol}(K_\delta) \geq \epsilon^n > \epsilon'$ , given a contradiction to the assertion in (ii).

By (20), the radius of the bounding ball can be chosen as  $H := 2b^L$ . Furthermore, the ellipsoid method works only with numbers having precision of  $O(N)$  bits. By Claim 2, the separation problem can be solved in time  $\text{poly}(n, U, D^k, \log \frac{1}{\delta})$ . □

**Remark 1** *By raising to inequalities (8) and (15) to power  $D$ , we obtain systems of polynomial inequalities. Khachiyan [Kha83, Kha84] gave a polynomial-time algorithm for (approximately) solving a system of convex polynomial inequalities. However, it is not possible to use this algorithm directly to solve the convex programs (6)-(9) and (13)-(16), since the polynomials obtained after raising inequalities (8) and (15) to power  $D$  are not necessarily convex. For instance, the function  $\sqrt{xy} - z$  is concave for  $x, y, z \in \mathbb{R}_+$ , while the function  $xy - z^2$  is not.*

<sup>2</sup>in case of  $K'_\delta$ , they are rational numbers of denominator at most  $n^{\Lambda D^2 U}$

<sup>3</sup>For this part of the proof we can consider the restriction of  $y$  to the set of positions reachable from  $v$  by one move. We can also replace parallel edges by one edge; if  $v$  is a position of chance then the transition probability of this edge is the sum of the transition probabilities of all corresponding parallel edges.

## 5 A Pseudo-polynomial algorithm for $k = O(1)$

Let  $\mathcal{G}$  be a BWR-game. Let  $t_{\max} := \max_{v \in V} \mu_{\mathcal{G}}(v)$  and  $t_{\min} := \min_{v \in V} \mu_{\mathcal{G}}(v)$ . Define the *top* and *bottom* classes of  $\mathcal{G}$  as  $\mathcal{T} := \{v \in V \mid \mu_{\mathcal{G}}(v) = t_{\max}\}$  and  $\mathcal{B} := \{v \in V \mid \mu_{\mathcal{G}}(v) = t_{\min}\}$ , respectively.

**Proposition 1** *Top and bottom classes necessarily satisfy the following properties.*

- (i) *There exists no arc  $(v, u) \in E$  such that  $v \in (V_W \cup V_R) \cap \mathcal{B}$ ,  $u \notin \mathcal{B}$ ;*
- (ii) *there exists no arc  $(v, u) \in E$  such that  $v \in (V_B \cup V_R) \cap \mathcal{T}$ ,  $u \notin \mathcal{T}$ ;*
- (iii) *there exists no arc  $(v, u) \in E$  such that  $v \in V_W \setminus \mathcal{T}$ ,  $u \in \mathcal{T}$ ;*
- (iv) *there exists no arc  $(v, u) \in E$  such that  $v \in V_B \setminus \mathcal{B}$ ,  $u \in \mathcal{B}$ ;*
- (v) *for every  $v \in V_W \cap \mathcal{T}$ , there exists an arc  $(v, u) \in E$  such that  $u \in \mathcal{T}$ ;*
- (vi) *for every  $v \in V_B \cap \mathcal{B}$ , there exists an arc  $(v, u) \in E$  such that  $u \in \mathcal{B}$ ;*
- (vii) *for every  $v \in (V_B \cup V_R) \setminus \mathcal{T}$ , there exists an arc  $(v, u) \in E$  such that  $u \notin \mathcal{T}$ ;*
- (viii) *for every  $v \in (V_W \cup V_R) \setminus \mathcal{B}$ , there exists an arc  $(v, u) \in E$  such that  $u \notin \mathcal{B}$ .*

**Proof** All claims follow from the existence of a canonical form for  $\mathcal{G}$ , by Theorem 2. Indeed, the existence of arcs forbidden by (i), (ii), (iii) and (iv), or the non-existence of arcs required by (v), (vi), (vii) and (viii) would violate the value equations (C1) of the canonical form.  $\square$

**Lemma 5** *Consider the convex program defined by (6)-(8) (resp., (13)-(15)). Then for  $t := t_{\max}$  (resp.,  $t := t_{\min}$ ), there is a feasible solution with  $y(v) \geq b^{-L}$  for all  $v \in \mathcal{T}$  (resp.,  $v \in \mathcal{B}$ ).*

**Proof** Consider the game  $\mathcal{G}[\mathcal{T}]$  (resp.,  $\mathcal{G}[\mathcal{B}]$ ) induced by the top class  $\mathcal{T}$  (resp., the bottom class  $\mathcal{B}$ ). Let  $x \in \mathbb{R}^{\mathcal{T}}$  (resp.,  $x \in \mathbb{R}^{\mathcal{B}}$ ) be the potential vector guaranteed by Theorem 2 for the game  $\mathcal{G}[\mathcal{T}]$  (resp.,  $\mathcal{G}[\mathcal{B}]$ ). Set  $y(v) := b^{-x(v)}$  for  $v \in \mathcal{T}$  (resp.,  $y(v) := b^{x(v)}$  for  $v \in \mathcal{B}$ ) and  $y(v) = 0$  for  $v \in V \setminus \mathcal{T}$  (resp.,  $v \in V \setminus \mathcal{B}$ ). Then  $y(v) \geq b^{-L}$  for all  $v \in \mathcal{T}$  (resp.,  $v \in \mathcal{B}$ ). It is easy to verify by Proposition 1 that the system is feasible. Indeed, (6) is satisfied for every position  $v \in V_W \cap \mathcal{T}$  (resp.,  $v \in V_B \cap \mathcal{B}$ ) by the definition of  $x$  and Fact 1:

$$t_{\max} \leq \max_{u|(v,u) \in E} (r(v,u) + x(v) - x(u)) \leq \log_b \sum_{u|(v,u) \in E} b^{r(v,u) + x(v) - x(u)}$$

$$(\text{resp., } -t_{\min} \leq -\min_{u|(v,u) \in E} (r(v,u) + x(v) - x(u)) \leq -\log_b \sum_{u|(v,u) \in E} b^{-r(v,u) - x(v) + x(u)}).$$

Moreover, for  $v \in (V_B \cup V_R) \cap \mathcal{T}$  (resp.,  $v \in (V_W \cup V_R) \cap \mathcal{B}$ ) we have (7) and (8) (resp., (14) and (15)) satisfied by the definition of  $x$  and Proposition 1-(ii) (resp., Proposition 1-(i)), while for  $v \in V \setminus \mathcal{T}$  (resp.,  $v \in V \setminus \mathcal{B}$ ) (6)-(8) (resp., (13)-(15)) are trivially satisfied.  $\square$

In the following we set  $\delta(t) := \frac{1}{2}b^{-t-L}(1 - \frac{1}{n})$ , where  $\varepsilon := \frac{1}{\sqrt{k}2^{k/2+1}D^{k+1}}$ . Note that  $b := n^{\frac{2}{\varepsilon}}$ .

**Lemma 6** *The values  $t_{\max}$  and  $t_{\min}$  can be found in time  $\text{poly}(n, U, D^k)$ .*

**Proof** We only show how to find  $t_{\max}$ ; in a similar fashion we can determine  $t_{\min}$ . We apply Lemma 4 in a binary search manner to check the feasibility of the system (13)-(16) for  $t \in [0, U]$  and  $\delta(t)$  as specified above. Note that, by Lemma 1,  $t_{\max} \in [0, U]$  can be written as a rational number with denominator at most  $\frac{1}{2\varepsilon}$ . So we may restrict our search steps to integer multiples of  $\frac{1}{2\varepsilon}$ . We stop the search when the length of the search interval becomes a constant multiple of  $\frac{1}{2\varepsilon}$ , and then apply linear search for the remaining small interval.



Suppose that the convex program (13)-(16) is infeasible. Then Theorem 2 implies that  $t_{\max} > t$ . On the other hand, if  $y \in \mathbb{R}^V$  is a  $\delta(t)$ -approximately feasible solution for (13)-(16), then as  $\delta(t) \leq \frac{1}{2}b^{-L}$ , the new vector  $y' := 2y$  satisfies  $y'(v) \in [b^{-L}, 2b^L + b^{-L}]$  for all  $v \in V$ . Also,  $y'$  satisfies (13)-(15) within an error of  $2\delta(t)$ , that is, the left-hand sides of (13)-(15), when  $y$  is replaced by  $y'$ , are at least  $b^{-t}y'(v) - 2\delta(t) = b^{-t}y'(v) - b^{-t-L}(1 - \frac{1}{n}) \geq b^{-t-\log_b n}y'(v)$ . Set  $x(v) := \log_b y'(v)$ . Then  $x$  satisfies (10)-(12) with  $t$  replaced by  $t + \log_b n$ . This in turn implies by Fact 1 that  $M[r_x] \leq (t + 2\log_b n)\mathbf{1} = (t + \varepsilon)\mathbf{1}$ . It follows then from Lemma 3 that  $t_{\max} \leq t + \varepsilon$ . Recall that we assume both  $t$  and  $t_{\max}$  are multiples of  $2\varepsilon$ ; hence,  $t_{\max} \leq t$ .

Since the number of binary search steps is at most  $\log \frac{U}{2\varepsilon} = O(k \log(UD))$  and each step requires time  $\text{poly}(n, U, \log b, \log \frac{1}{\delta}) = \text{poly}(n, U, k)$ , the bound on the running time follows.  $\square$

**Lemma 7** *We can find the top class  $\mathcal{T}$  (resp., bottom class  $\mathcal{B}$ ) in time  $\text{poly}(n, U, D^k)$ .*

**Proof** We can check if a vertex  $w \in V$  belongs to the top class (resp., bottom class) as follows. We write the convex program (6)-(8) (resp., (13)-(15)) with  $t := t_{\max}$  (resp.,  $t := t_{\min}$ ) and with the additional constraint that  $y(w) \geq b^{-L}$  and  $y(v) \geq 0$  for all  $v \in V \setminus \{w\}$ . Then we check the feasibility of this system. If the system is infeasible then we know by Lemma 5 that  $w \notin \mathcal{T}$  (resp.,  $w \notin \mathcal{B}$ ).

Suppose, on the other hand, that  $y \in \mathbb{R}^V$  is a  $\delta(t_{\max})$ -approximately (resp.,  $\delta(t_{\min})$ -approximately) feasible solution for (6)-(8) (resp., (13)-(15)). Then as in the proof of Lemma 6, the new vector  $y' := 2y$  satisfies  $y'(w) \geq b^{-L} > 0$ , and the left-hand sides of (6)-(8) (resp., (13)-(15)), when  $y$  is replaced by  $y'$ , are at least  $b^{t-\log_b n}y'(w)$  (resp.,  $b^{-t-\log_b n}y'(w)$ ).

Now we claim that  $V^+ := \{v \in V : y(v) > 0\} \subseteq \mathcal{T}$  (resp.,  $V^+ := \{v \in V : y(v) > 0\} \subseteq \mathcal{B}$ ), which would in turn imply that  $w \in \mathcal{T}$  (resp.,  $w \in \mathcal{B}$ ). Indeed, constraints (6)-(8) (resp., (13)-(15)), applied to  $y$  replaced by  $y'$ , imply that (i) if  $v \in V_W \cap V^+$  then there exists an arc  $(v, u) \in E$  such that  $u \in V^+$ ; (ii) if  $v \in (V_B \cup V_R) \cap V^+$  then all arcs  $(v, u) \in E$  must that  $u \in V^+$  (resp., (i) if  $v \in V_B \cap V^+$  then there exists an arc  $(v, u) \in E$  such that  $u \in V^+$ ; (ii) if  $v \in (V_W \cup V_R) \cap V^+$  then all arcs  $(v, u) \in E$  must that  $u \in V^+$ ). These imply that the game induced by  $V^+$  is well-defined and, by Lemma 2 (resp., Lemma 3), all its positions have value at least  $t_{\max}$  (resp., at most  $t_{\min}$ ). The lemma follows.  $\square$

Finally, given the top and bottom classes, we can find an optimal pair of strategies in the games induced by  $\mathcal{T}$  and  $\mathcal{B}$ , as stated in the next lemma.

**Lemma 8** *We can find optimal pairs of strategies in the games induced by the top class  $\mathcal{T}$  and bottom class  $\mathcal{B}$  in time  $\text{poly}(n, U, D^k)$ .*

**Proof** We prove the lemma only for  $\mathcal{T}$ ; the proof for  $\mathcal{B}$  can be done similarly. We solve two (feasible) systems,  $\mathcal{S}_1$  defined by (6)-(9) on  $\mathcal{G}[\mathcal{T}]$  and  $\mathcal{S}_2$  defined by (13)-(15) on  $\mathcal{G}[\mathcal{T}]$ , with  $t := t_{\max}$  to within an accuracy of  $\delta(t_{\max})$ . Let  $y^1, y^2 \in \mathbb{R}^{\mathcal{T}}$  be the  $\delta(t_{\max})$ -approximate solutions to  $\mathcal{S}_1$  and  $\mathcal{S}_2$ , respectively. By the same arguments as in Lemma 6, the corresponding potential vectors  $x^1, x^2$  (defined by  $x^1(v) := -\log_b(2y^1(v))$  and  $x^2(v) := \log_b(2y^2(v))$ ) ensure that  $M[r_{x^1}] \geq (t_{\max} - \varepsilon)\mathbf{1}$  and  $M[r_{x^2}] \leq (t_{\max} + \varepsilon)\mathbf{1}$ . Since  $\varepsilon$  is sufficiently small, by Lemmas 2 and 3, the locally optimal strategies defined by the operator  $M$  with respect to  $x^1$  and  $x^2$  give optimal strategies for MAX and MIN in  $\mathcal{G}[\mathcal{T}]$ , respectively.  $\square$

Finally, we obtain Theorem 1 by combining the above lemmas with the algorithm in [BEGM13b, BEGM15].

## References

- [AM09] D. Andersson and P. B. Miltersen. The complexity of solving stochastic games on graphs. In *ISAAC*, pages 112–121, 2009.

- [BEGM10] E. Boros, K. Elbassioni, V. Gurvich, and K. Makino. A pumping algorithm for ergodic stochastic mean payoff games with perfect information. In *Proc. 14th IPCO*, pages 341–354, 2010.
- [BEGM13a] E. Boros, K. Elbassioni, V. Gurvich, and K. Makino. On canonical forms for zero-sum stochastic mean payoff games. *Dynamic Games and Applications*, 3(2):128–161, 2013.
- [BEGM13b] Endre Boros, Khaled M. Elbassioni, Vladimir Gurvich, and Kazuhisa Makino. A pseudo-polynomial algorithm for mean payoff stochastic games with perfect information and a few random positions. In *Automata, Languages, and Programming - 40th International Colloquium, ICALP 2013, Riga, Latvia, July 8-12, 2013, Proceedings, Part I*, pages 220–231, 2013.
- [BEGM15] E. Boros, K. Elbassioni, V. Gurvich, and K. Makino. A pseudo-polynomial algorithm for mean payoff stochastic games with perfect information and few random positions. *CoRR*, abs/1508.03431, 2015.
- [BV01a] E. Beffara and S. Vorobyov. Adapting gurvich-karzanov-khachiyan’s algorithm for parity games: Implementation and experimentation. Technical Report 2001-020, Department of Information Technology, Uppsala University, available at: <https://www.it.uu.se/research/reports/#2001>, 2001.
- [BV01b] E. Beffara and S. Vorobyov. Is randomized gurvich-karzanov-khachiyan’s algorithm for parity games polynomial? Technical Report 2001-025, Department of Information Technology, Uppsala University, available at: <https://www.it.uu.se/research/reports/#2001>, 2001.
- [BV05] H. Björklund and S. Vorobyov. Combinatorial structure and randomized sub-exponential algorithm for infinite games. *Theoretical Computer Science*, 349(3):347–360, 2005.
- [BV07] H. Björklund and S. Vorobyov. A combinatorial strongly sub-exponential strategy improvement algorithm for mean payoff games. *Discrete Applied Mathematics*, 155(2):210–229, 2007.
- [CH08] K. Chatterjee and T. A. Henzinger. Reduction of stochastic parity to stochastic mean-payoff games. *Inf. Process. Lett.*, 106(1):1–7, 2008.
- [CHKN14] K. Chatterjee, M. Henzinger, S. Krininger, and D. Nanongkai. Polynomial-time algorithms for energy games with special weight structures. *Algorithmica*, 70(3):457–492, 2014.
- [CJH04] K. Chatterjee, M. Jurdziński, and T. A. Henzinger. Quantitative stochastic parity games. In *SODA ’04*, pages 121–130, Philadelphia, PA, USA, 2004. Society for Industrial and Applied Mathematics.
- [EM79] A. Eherenfeucht and J. Mycielski. Positional strategies for mean payoff games. *International Journal of Game Theory*, 8:109–113, 1979.
- [Gil57] D. Gillette. Stochastic games with zero stop probabilities. In A.W. Tucker M. Dresher and P. Wolfe, editors, *Contribution to the Theory of Games III, in Annals of Mathematics Studies*, volume 39, pages 179–187. Princeton University Press, 1957.
- [GKK88] V. Gurvich, A. Karzanov, and L. Khachiyan. Cyclic games and an algorithm to find minimax cycle means in directed graphs. *USSR Computational Mathematics and Mathematical Physics*, 28:85–91, 1988.
- [GLS88] M. Grötschel, L. Lovász, and A. Schrijver. *Geometric Algorithms and Combinatorial Optimization*. Springer, New York, 1988.

- [Hal07] N. Halman. Simple stochastic games, parity games, mean payoff games and discounted payoff games are all LP-type problems. *Algorithmica*, 49(1):37–50, 2007.
- [HBV04] S. Sandberg H. Björklund and S. Vorobyov. A combinatorial strongly sub-exponential strategy improvement algorithm for mean payoff games. DIMACS Technical Report 2004-05, DIMACS, Rutgers University, 2004.
- [JPZ06] M. Jurdziński, M. Paterson, and U. Zwick. A deterministic subexponential algorithm for solving parity games. In *SODA '06*, pages 117–123, New York, NY, USA, 2006. ACM.
- [Jur98] M. Jurdziński. Deciding the winner in parity games is in  $UP \cap co-UP$ . *Inf. Process. Lett.*, 68(3):119–124, 1998.
- [Jur00] M. Jurdziński. *Games for Verification: Algorithmic Issues*. PhD thesis, Department of Computer Science, University of Aarhus, Denmark, 2000.
- [Kar78] R. M. Karp. A characterization of the minimum cycle mean in a digraph. *Discrete Math.*, 23:309–311, 1978.
- [Kha80] L. Khachiyan. Polynomial algorithms in linear programming. *U.S.S.R. Comput. Math. Math. Phys.*, 20:53–72, 1980.
- [Kha83] L. Khachiyan. Convexity and complexity in polynomial programming. In *Proceedings of the International Congress of Mathematicians, Warsaw*, page 15691577, 1983.
- [Kha84] L. Khachiyan. *Complexity of convex problems of real and integer programming*. PhD thesis, The second thesis, manuscript, Computer Center of the USSR Academy of Sciences (in Russian), Moscow, 1984.
- [KL93] A. V. Karzanov and V. N. Lebedev. Cyclical games with prohibition. *Mathematical Programming*, 60:277–293, 1993.
- [Lit96] M. L. Littman. *Algorithm for sequential decision making, CS-96-09*. PhD thesis, Dept. of Computer Science, Brown Univ., USA, 1996.
- [LL69] T. M. Liggett and S. A. Lippman. Stochastic games with perfect information and time-average payoff. *SIAM Review*, 4:604–607, 1969.
- [MO70] H. Mine and S. Osaki. *Markovian decision process*. American Elsevier Publishing Co., New York, 1970.
- [Mou76a] H. Moulin. Extension of two person zero sum games. *Journal of Mathematical Analysis and Application*, 5(2):490–507, 1976.
- [Mou76b] H. Moulin. Prolongement des jeux à deux joueurs de somme nulle. *Bull. Soc. Math. France, Memoire*, 45, 1976.
- [Pis99] N. N. Pisaruk. Mean cost cyclical games. *Mathematics of Operations Research*, 24(4):817–828, 1999.
- [RF91] T. E. S. Raghavan and J. A. Filar. Algorithms for stochastic games: A survey. *Mathematical Methods of Operations Research*, 35(6):437–472, 1991.
- [Sch09] Sven Schewe. From parity and payoff games to linear programming. In *Mathematical Foundations of Computer Science 2009, 34th International Symposium, MFCS 2009, Novy Smokovec, High Tatras, Slovakia, August 24-28, 2009. Proceedings*, pages 675–686, 2009.
- [Vor08] S. Vorobyov. Cyclic games and linear programming. *Discrete Applied Mathematics, Special volume in Memory of Leonid Khachiyan (1952 - 2005)*, 156(11):2195–2231, 2008.

- [ZP96] U. Zwick and M. Paterson. The complexity of mean payoff games on graphs. *Theoretical Computer Science*, 158(1-2):343–359, 1996.